

Mon. Not. R. Astron. Soc. **000**, 000,000 (1999)

Noise properties of gravitational lens mass reconstruction

Ludovic Van Waerbeke

Canadian Institute for Theoretical Astrophysics, 60 St George Str., Toronto, M5S 3H8, Canada

10 September 1999

ABSTRACT

Gravitational lensing is potentially able to observe mass-selected halos, and to measure the projected cluster mass function. An optimal mass-selection requires a quantitative understanding of the noise behavior in mass maps. This paper is an analysis of the noise properties in mass maps reconstructed using a maximum likelihood method.

The first part of this work is the derivation of the noise power spectrum and the mass error bars as a straightforward extension of the Kaiser & Squires (1993) algorithm to the case of a correlated noise. A very good agreement is found between these calculations and the noise properties observed in the mass reconstructions limited to non-critical cluster of galaxies. It demonstrates that Kaiser & Squires (1993) and maximum likelihood methods have similar noise properties and that the weak lensing approximation is valid for describing it.

In a second stage I show that the statistic of peaks in the noise follows accurately the peak statistics of a two-dimensional Gaussian random field (using the BBKS technics) if the smoothing aperture contains enough galaxies. This analysis provides a full procedure to derive the significance of any convergence peak as a function of its amplitude and its profile.

It is demonstrated that, to a very good approximation, a mass map is the sum of the lensing signal plus a 2D gaussian random noise, which means that a detailed quantitative analysis of the structures in mass maps can be done. A straightforward application is the measurement of the projected mass function in wide field lensing surveys, down to small mass over-densities which are individually undetectable.

Key words: cosmology: gravitational lensing, dark matter - galaxies: clusters

1 INTRODUCTION

Non-parametric mass reconstruction from gravitational shear is now a very well understood problem, and was widely refined over the last few years. It is no longer a technical issue to reconstruct clusters (see Squires & Kaiser 1996 for a detailed review) as well as large scale structures (Van Waerbeke et al. 1999). Although the computational speed of mass reconstruction might be too low for large scale surveys (Lombardi & Bertin 1999), this is not a major concern yet according to the actual size of the lensing surveys.

Naturally a next step in gravitational lensing is the search for a better use of the information present in the mass maps. Assuming that all the problems concerning the residual of Point Spread Function corrections are or will be solved, there is useful information hidden in the noise of mass maps that can be extracted. An important application of lensing is the possibility to built mass-selected samples (Reblinsky & Bartelmann, 1999) as a probe of cosmology (Oukbir & Blanchard 1997, Reblinsky et al. 1999). In par-

ticular a joint analysis of clusters of galaxies from weak lensing, Sunyaev-Zeldovich, X-rays and galaxy catalogues would be valuable. A related issue is the possible existence of *dark lenses* (Hattori 1997), and already, the detection of a mass clump not obviously associated with light (Erben et al. 1999) raised the question of quantifying accurately the significance of halo detections.

The understanding of the noise properties in mass maps is therefore necessary for anyone who wants to measure the mass distribution from gravitational lensing in order to constrain cosmological models. This point was understood by Schneider (1996) who constructed the so-called S-statistic, aimed to quantify the significance of mass peaks. Using bootstrapping of the data, he derives the frequency of observed mass peaks. The S-statistic is similar to the the signal-to-noise derived in Kaiser & Squires (1993) (and confirmed later in numerical simulations by Wilson et al. 1996) but it has the interesting feature that works directly with the ellipticity of the galaxies (no mass reconstruction is required),

and the smoothing filter can be optimized in order to increase the signal-to-noise.

So far, the noise analysis and the statistical significance of mass reconstructions have not been studied as much as the mass reconstruction methods themselves. We have only some global clues of the behavior of the noise, which above all, seem to depend on the mass reconstruction method (Squires & Kaiser 1996). Seitz & Schneider (1996) identified several characteristics of the noise in their numerical simulations: due to the smoothing, the power spectrum is not flat and driven by the intrinsic ellipticity of the galaxies, their number density and the smoothing window scale. Lombardi & Bertin (1998a, 1998b) gave the first analytical estimate of the variance of the total mass of a reconstructed cluster, taking into account the noise correlation induced by the smoothing window. They have found differences with the numerical results of Seitz & Schneider (1996) that they attributed to *technical* differences. They worked with shear line-integral reconstruction methods, and found that the noise is minimized when curl-free kernels are used. As discussed in Seitz & Schneider (1996) the curl-free kernels have a noise amplitude similar to the Kaiser & Squires (1993) (hereafter KS) reconstruction, but the line-integral reconstruction is a non-local process, which complicates the noise analysis (Lombardi & Bertin (1998a, 1998b)). On the other hand, the Maximum Likelihood (hereafter ML) algorithm (Bartelmann et al. 1996) performs a local reconstruction, but its noise properties can hardly be studied analytically. In fact, the KS scheme seems to be the easiest approach to an analytical description of the noise.

It has been demonstrated in the case of lensing by large scale structures (see the simulations performed in Van Waerbeke et al. 1999) that the noise in ML reconstructions is accurately described by the weak lensing approximation (and therefore by the KS equations). Continuing this work, this paper is a study of the noise properties of ML methods in the case of non-critical clusters of galaxies with correlated noise, when the smoothing is taken into account.

This paper shows that the noise in ML reconstruction agrees very well with the analytical prediction made from the weak lensing approximation. It demonstrates that the mass variance derived in Lombardi & Bertin (1998b) is valid, and, to a larger extent, that the noise is uncorrelated with the signal even for non-critical clusters of galaxies. This is also consistent with the empirical analysis done in Seitz & Schneider (1996). Assuming that the smoothed intrinsic ellipticities are described by the statistics of a two-dimensional Gaussian random field, the BBKS technique (Bardeen et al. 1986) is used in order to quantify the significance of the overdensities in ML mass reconstructions. This analysis provides a quantitative way to identify the structures in mass maps and to extract information from the noise: this result is verified on numerical simulations, which shows that to a very good approximation a mass map is the sum of a completely known Gaussian field plus the lensing signal.

Section 2 gives some general considerations about lensing theory used in this work. The noise model is developed in Section 3, where each noise contribution is identified. An estimator for the variance of the mass in a given area is calculated. In Section 4, the noise model and numerical simulations of cluster lensing are compared. In Section 5 the significance level of a reconstructed mass halo based on a

2D Gaussian field statistic is derived, which is then checked against numerical simulations.

2 GENERALITIES

We consider a lens of surface mass density $\Sigma(\boldsymbol{\theta})$ located at a redshift z_l along the line of sight $\boldsymbol{\theta}$, and a single source plane is put at a redshift z_s . At first approximation, the Universe is equivalent to a single lens, so there is no loss of generality here to consider a single lens. The projected gravitational potential Φ and the surface mass density of the lens are given by the 2-dimensional Poisson equation

$$\Delta\Phi(\boldsymbol{\theta}) = 4\pi G\Sigma(\boldsymbol{\theta}). \quad (1)$$

The critical density (for which the lens produces multiple images) is given by

$$\Sigma_{crit} = \frac{c^2}{4\pi G} \frac{\mathcal{D}_{os}}{\mathcal{D}_{ol}\mathcal{D}_{ls}}, \quad (2)$$

where \mathcal{D}_{ij} is the angular diameter distance between the redshifts z_i and z_j . For convenience, the lensing potential is defined as $\phi = \Phi/(2\pi G\Sigma_{crit})$. The ratio

$$\kappa(\boldsymbol{\theta}) = \frac{\Sigma(\boldsymbol{\theta})}{\Sigma_{crit}}, \quad (3)$$

is the convergence of the lens, which describes the isotropic deformation of a light beam coming from the sources. The light beam is also sheared, which modifies the shape of the source galaxies. The shape of a galaxy is described by the second moments of its centered surface brightness profile

$$Q_{ij} = \int d\boldsymbol{\theta} \theta_i \theta_j S(\boldsymbol{\theta}) / \int d\boldsymbol{\theta} S(\boldsymbol{\theta}). \quad (4)$$

The source and image shape matrices are related by the equation

$$Q_{ij}^{(I)} = \mathcal{A} Q_{ij}^{(S)} \mathcal{A}, \quad (5)$$

where \mathcal{A} is the magnification matrix which depends on the convergence κ and the shear $\boldsymbol{\gamma}$ of the lens,

$$\mathcal{A} = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix}. \quad (6)$$

κ and $\boldsymbol{\gamma}$ are fully determined by the second derivatives of the lensing potential via

$$\kappa = \Delta\phi/2; \quad \gamma_1 = (\phi_{11} - \phi_{22})/2; \quad \gamma_2 = \phi_{12}. \quad (7)$$

The ellipticity is defined in the complex plane as

$$\epsilon = \frac{1-r}{1+r} e^{2i\vartheta}, \quad (8)$$

where r is the axis ratio of a galaxy and ϑ the orientation. Eq.(5) written in terms of source and image ellipticities is

$$\epsilon^{(I)} = \frac{\epsilon^{(S)} + \mathbf{g}}{1 + \mathbf{g}^* \epsilon^{(S)}}, \quad (9)$$

where $\mathbf{g} = \boldsymbol{\gamma}/(1 - \kappa)$ is the complex reduced shear. The observable is $\boldsymbol{\delta} = 2\mathbf{g}/(1 + |\mathbf{g}|^2)$, but this paper is restricted to non-critical lenses, for which $|\boldsymbol{\delta}| < 1$, so that \mathbf{g} is an observable as well. In that case \mathbf{g} can be immediately read off from the galaxy ellipticities; as shown by Schramm &

Kayser (1995) and Schneider & Seitz (1995) the mean observed ellipticities of the galaxies is an unbiased estimate of the reduced shear:

$$\langle \epsilon^{(I)} \rangle = \mathbf{g}. \quad (10)$$

The mass reconstructed from shear measurement alone is not fully constrained because any transformation of the lensing potential of the form

$$\phi \rightarrow \lambda\phi + \frac{(1-\lambda)}{2}|\theta|^2 \quad (11)$$

leaves the reduced shear \mathbf{g} unchanged. This is the so-called mass-sheet degeneracy: the mass is determined up to an additive constant and an appropriate scaling (Eq.(11) implies $\kappa \rightarrow 1 - \lambda + \lambda\kappa$).

3 NOISE CORRELATION AND MASS ERROR BAR

3.1 Weak lensing approximation

As already mentioned, the weak lensing approximation linearises the lensing equation and dramatically simplifies the noise analysis. The weak lensing limit of Eq.(9) is

$$\epsilon^{(I)} \simeq \gamma + \epsilon^{(S)}. \quad (12)$$

Thus $\langle \epsilon^{(I)} \rangle = \gamma$ becomes an unbiased estimate of the shear. In the Fourier space Eq.(7) takes the form (Kaiser 1998):

$$\tilde{\kappa}(\mathbf{k}) = c_\alpha(k)\tilde{\gamma}_\alpha(\mathbf{k}), \quad (13)$$

where there is an implicit summation over $\alpha = (1, 2)$. $c_\alpha = (\cos(2\varphi), \sin(2\varphi))$, and $\mathbf{k} = k(\cos\varphi, \sin\varphi)$. An unbiased (but extremely noisy) estimate of the convergence κ is $\kappa_n(\theta) = \kappa(\theta) + n(\theta)$, where $n(\theta)$ is the noise part, which depends on the intrinsic ellipticities of the galaxies and their spatial distribution. The Fourier transform of κ_n is

$$\tilde{\kappa}_n(\mathbf{k}) = c_\alpha \tilde{\epsilon}_\alpha^{(I)}(\mathbf{k}) = \tilde{\kappa}(\mathbf{k}) + c_\alpha \tilde{\epsilon}_\alpha^{(S)}(\mathbf{k}). \quad (14)$$

In order to avoid the infinite variance problem (caused by the discreteness distribution of the galaxies), and to have a mass map with a reasonable amplitude of noise, the convergence is reconstructed from the smoothed galaxy ellipticities, so smoothed fields are used in the following. Uppercase letters refer to smoothed fields while lowercase letters refer to unsmoothed fields (e.g. K for the smoothed κ , $\mathbf{\Gamma}$ for γ and $\mathbf{\Sigma}$ for ϵ). The smoothing function is denoted by $W(|\theta|)$, the total number of galaxies in the field by N_g , and the total field area by \mathcal{S} . Galaxies are at positions θ_i . The reconstructed mass map is now the smoothed convergence with an unbiased estimate given by

$$K(\theta) = \frac{\mathcal{S}}{N_g} \sum_{i=1}^{N_g} W(|\theta - \theta_i|) \kappa(\theta_i) \quad (15)$$

The number density of galaxies is $n_g = N_g/\mathcal{S}$. It is worth noting that Eq.(13) implies that the reconstructed convergence of a smoothed shear field is identical to the initial smoothed convergence. In other words, smoothing and reconstruction are commutative processes in the frame of the weak lensing approximation.

3.2 Noise correlation function

The reconstructed convergence $K_N(\theta)$ is the sum of the true convergence $K(\theta)$ and a correlated noise $N(\theta)$. If $\mathbf{\Sigma}^{(I)}$ and $\mathbf{\Sigma}^{(S)}$ are the smoothed $\epsilon^{(I)}$ and $\epsilon^{(S)}$ then

$$K_N(\theta) = K(\theta) + N(\theta) = \int d\mathbf{k} e^{-i\mathbf{k}\cdot\theta} c_\alpha(k) \tilde{\Sigma}_\alpha^{(I)}(\mathbf{k}). \quad (16)$$

If the smoothing length is small enough (arcmin scale) the signal is approximately constant within the aperture. This is certainly not true for cluster mass profiles, but it should still give a reasonable description of the noise. It implies that the sum over the galaxy ellipticities within the aperture is equal to the smoothed shear plus the intrinsic ellipticity contribution (which is Eq.(12) for the smoothed fields):

$$\mathbf{\Sigma}^{(I)}(\theta) \simeq \mathbf{\Gamma}(\theta) + \frac{1}{n_g} \sum_{i=1}^{N_g} W(\theta - \theta_i) \epsilon^{(S)}(\theta_i). \quad (17)$$

Inserting this into Eq.(16) gives the analytic expression of the noise term:

$$K_N(\theta) - K(\theta) \simeq \frac{1}{n_g} \sum_{i=1}^{N_g} \int d\mathbf{k} \tilde{W}(\mathbf{k}) e^{-i\mathbf{k}\cdot(\theta - \theta_i)} c_\alpha(k) \epsilon_\alpha(\theta_i), \quad (18)$$

where $\tilde{W}(\mathbf{k})$ is the Fourier transform of the window function. The ensemble average of the square of Eq.(18) is the noise variance $\langle N^2(\theta) \rangle$. It is calculated by averaging over the intrinsic ellipticities and the galaxy positions. Let A be the operator which averages over the intrinsic ellipticities and P the operator which averages over the galaxy positions. The former is defined as

$$A \left[\epsilon_i^{(S)}(\theta) \epsilon_j^{(S)}(\theta') \right] = \frac{\sigma_\epsilon^2}{2} \delta_{ij}^K \delta_D(\theta - \theta'). \quad (19)$$

δ_{ij}^K is the Kronecker symbol, and σ_ϵ the intrinsic ellipticity dispersion. This is a consequence of the fact that the galaxy ellipticities are assumed to be spatially uncorrelated. The galaxy position average operator P depends on the angular clustering of the galaxies within the aperture (Peebles 1980):

$$P = \int \frac{d^2\theta_1 d^2\theta_2}{\mathcal{S}^2} (1 + \omega(|\theta_2 - \theta_1|) + \dots) \quad (20)$$

where \mathcal{S} is total field area, and $\omega(|\theta_2 - \theta_1|)$ is the angular two point correlation function of the source galaxies. This clustering term is generally ignored, and accounts for possible increase of the variance of $N(\theta)$ in the case of non-uniform sampling of the source galaxies within the aperture. Its contribution can be potentially important if the source galaxies belongs to a thin redshift slice (as the angular clustering approaches the spatial clustering). Connolly et al. (1998) measured on the Hubble Deep Field the angular correlation function of galaxies in the redshift interval $[1, 1.2]$, which corresponds to the most likely source redshift for weak lensing surveys. They found an angular correlation length of 3 arcsec, and therefore negligible for our purpose. In practice the redshift slice of the lensed galaxies is much larger $[0.5, 1.2]$, which makes the contribution of the angular clustering totally negligible (Le Fèvre et al. 1996). As a consequence, the ensemble average of the local variance of the convergence is simply given by:

$$P \left[A \left[N^2(\theta) \right] \right] = \frac{\sigma_\epsilon^2}{2n_g} \int d\mathbf{k} |\tilde{W}(\mathbf{k})|^2, \quad (21)$$

as derived in Kaiser & Squires 1993. The noise correlation function is then

$$P[A[N(\boldsymbol{\theta})N(\boldsymbol{\theta}')]] = \frac{\sigma_\epsilon^2}{2n_g} \int d\mathbf{k} e^{i\mathbf{k}\cdot(\boldsymbol{\theta}'-\boldsymbol{\theta})} |\tilde{W}(\mathbf{k})|^2. \quad (22)$$

3.3 Average mass error bar

The average mass error bar ΔM in an aperture (the total mass of a cluster for instance, or the mass of a sub-structure) is now calculated. Provided that the redshifts of the background galaxies are all the same, the projected mass $M(\boldsymbol{\theta})$ in the direction $\boldsymbol{\theta}$ is $M(\boldsymbol{\theta}) = \Sigma_{crit}(z_l, z_s) K(\boldsymbol{\theta})$, where $\Sigma_{crit}(z_l, z_s)$ is the critical surface mass density of the lens which depends on the lens configuration (Eq.(2)). If the background galaxies were distributed in redshift, the reduced shear becomes redshift dependent but it can be easily accounted for (Seitz & Schneider 1997), and with no loss of generality the analysis is restricted to a single source redshift.

The variance of the total projected mass estimated within an area \mathcal{S} is by definition

$$\Delta M^2 = \Sigma_{crit}^2 \int_{\mathcal{S}} \int_{\mathcal{S}} d\boldsymbol{\theta} d\boldsymbol{\theta}' P[A[N(\boldsymbol{\theta})N(\boldsymbol{\theta}')]]. \quad (23)$$

According to Eq.(22) the variance is finally

$$\Delta M^2 = \Sigma_{crit}^2 \frac{\sigma_\epsilon^2}{2} \int_{\mathcal{S}} \frac{d\boldsymbol{\theta}}{n_g} \int_{\mathcal{S}} d\mathbf{r} \int d\mathbf{k} e^{i\mathbf{k}\cdot\mathbf{r}} |\tilde{W}(\mathbf{k})|^2. \quad (24)$$

3.4 Gaussian smoothing

The Gaussian window has a width σ , and is normalized:

$$W(|\boldsymbol{\theta}|) = \frac{1}{\pi\sigma^2} \exp\left(-\frac{|\boldsymbol{\theta}|^2}{\sigma^2}\right). \quad (25)$$

Inserting this into Eq.(22) gives the noise correlation function

$$\langle N(\boldsymbol{\theta})N(\boldsymbol{\theta}') \rangle = \frac{\sigma_\epsilon^2}{2} \frac{1}{2\pi\sigma^2 n_g} \exp\left(-\frac{|\boldsymbol{\theta}' - \boldsymbol{\theta}|^2}{2\sigma^2}\right). \quad (26)$$

The variance defined by Eq.(23) for a square area $\mathcal{S} = [2a, 2a]$ is then straightforward to calculate:

$$\Delta M^2 = 4\sigma^4 \left(\frac{\sigma_\epsilon^2}{4\pi\sigma^2 n_g} \right) \Sigma_{crit}^2 \times \left[1 - \exp\left(-2\frac{a^2}{\sigma^2}\right) - \sqrt{2\pi}\frac{a}{\sigma} \operatorname{erf}\left(\sqrt{2}\frac{a}{\sigma}\right) \right]^2. \quad (27)$$

3.5 Top-hat smoothing

The normalised top-hat of radius σ is

$$W(|\boldsymbol{\theta}|) = \begin{cases} \frac{1}{\pi\sigma^2} & \text{if } |\boldsymbol{\theta}| < \sigma \\ 0 & \text{otherwise} \end{cases} \quad (28)$$

The noise correlation function is then

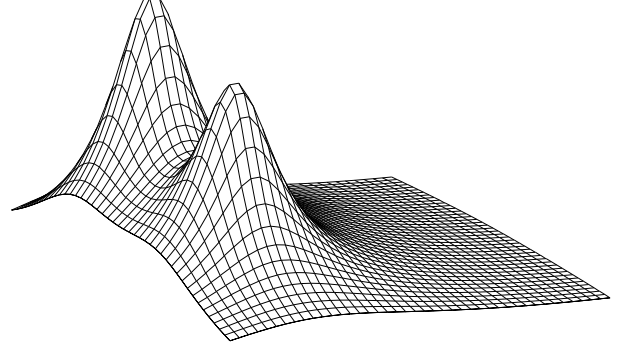


Figure 1. The cluster mass model used in this work sampled on a 42×42 pixels grid, with a maximum convergence $\kappa_{max} = 0.3$. The lens, defined by Eq.(31) has a redshift $z = 0.2$, and an Einstein-de-Sitter Universe is assumed. This is the mass distribution caused by two truncated isothermal potential with a core radius of 5 pixels.

$$\begin{aligned} \langle N(\boldsymbol{\theta})N(\boldsymbol{\theta}') \rangle &= \frac{\sigma_\epsilon^2}{2} \frac{1}{\pi\sigma^2 n_g} \frac{s(|\boldsymbol{\theta}' - \boldsymbol{\theta}|)}{\pi\sigma^2} \text{ if } |\boldsymbol{\theta}' - \boldsymbol{\theta}| < 2\sigma \\ &= 0 \text{ otherwise} \end{aligned} \quad (29)$$

where $s(|\boldsymbol{\theta}' - \boldsymbol{\theta}|)$ is the overlapping area of two identical circles of radius σ distant by $d = |\boldsymbol{\theta}' - \boldsymbol{\theta}|$

$$s(d) = \sigma^2 \left[2\arccos\left(\frac{d}{2\sigma}\right) - \frac{d}{\sigma} \sqrt{1 - \frac{d^2}{4\sigma^2}} \right], \quad (30)$$

4 CLUSTER LENS SIMULATIONS

4.1 Description

The description of the noise given above is now compared to mass reconstruction simulations using the lens comparable in lensing strength to a non-critical cluster of galaxies ($\kappa_{max} = 0.3$). The lens potential ϕ is the sum of two non-singular isothermal spheres with a core radius R_c , located at $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$:

$$\begin{aligned} \psi(\boldsymbol{\theta}) &= \frac{4\pi\sigma_\infty^2}{c^2} \frac{\mathcal{D}_{ol}\mathcal{D}_{ls}}{\mathcal{D}_{os}} R_c^2 \left(\sqrt{1 + \frac{|\boldsymbol{\theta} - \boldsymbol{\theta}_1|^2}{R_c^2}} \right. \\ &\quad \left. + 0.8 \sqrt{1 + \frac{|\boldsymbol{\theta} - \boldsymbol{\theta}_2|^2}{R_c^2}} \right). \end{aligned} \quad (31)$$

The velocity dispersion is $\sigma_\infty = 900$ km/s, $R_c = 5$, and we assume a lens redshift $z_l = 0.2$. The distances are calculated in an Einstein-de-Sitter Universe. Figure 1 shows the convergence $\kappa = \Delta\psi$ of the lens. The shear caused by the lens is applied to a background galaxy population at redshift $z_s = 1$, with an ellipticity distribution given by

$$p(\epsilon^{(S)}) \propto \exp\left[-\left(\epsilon^{(S)}/\sigma_\epsilon\right)^2\right], \quad (32)$$

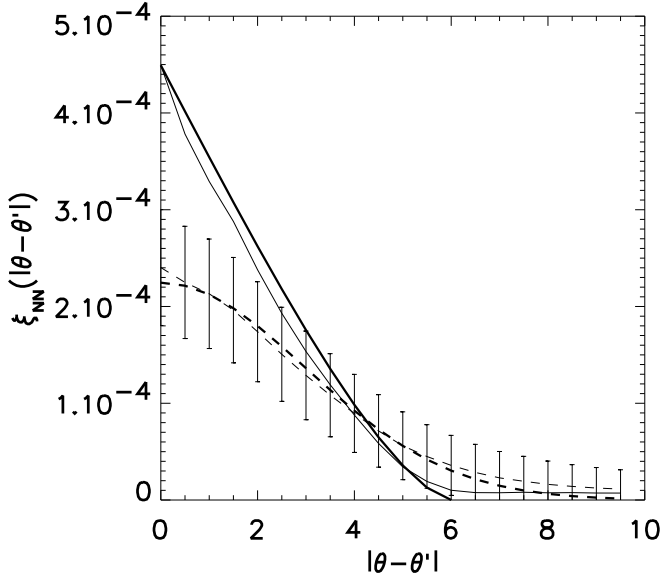


Figure 2. The noise correlation function measured in the simulations (thin lines), compared to the model (thick lines). The dashed lines correspond to the Gaussian smoothing and the solid lines for top-hat smoothing. Error bars are calculated from 50 different simulation realizations, they are displayed for the gaussian smoothing only, but they are similar for the top-hat case.

with $\sigma_\epsilon = 0.12$ (corresponding to a typical axis ratio of 0.8). 1000 background galaxies are randomly placed on a 42×42 pixels grid. An *observed* simulated shear map is then calculated, and the lens mass distribution is reconstructed with a maximum-likelihood algorithm (Bartelmann et al. 1996). The maximum likelihood reconstruction minimizes the χ^2 function defined as

$$\chi^2 = \sum_{i,j} |\mathbf{g}(\phi_{guess}) - \mathbf{g}_{obs}|^2, \quad (33)$$

where ϕ_{guess} is a guessed lensing potential. The reduced shear $\mathbf{g}(\phi_{guess})$ is calculated from ϕ_{guess} , and \mathbf{g}_{obs} is the observed reduced shear. The reduced shear is calculated from the finite difference scheme described in Van Waerbeke et al. (1999). The comparisons are done between the reconstructed mass map and the initial mass smoothed with the same window used to smooth the shear for the reconstructions (mass reconstruction and smoothing should commute according to the noise model). For each type of smoothing window (top-hat and Gaussian) 50 simulations are performed. The widths of the smoothing windows are $\sigma = 3$ pixels, therefore there is a mean number of 16 galaxies per aperture. An edge of width 3 pixels is systematically removed from the analysis because of its intrinsic higher noise due to the change of the finite difference scheme at the boundaries of the grid and the smaller number of galaxies in the aperture. The calibration constant λ (see Section 4.2) is calculated by forcing the total mass of each reconstructed map to be equal to the total mass of the smoothed mass distribution of the lens. Therefore the variance of the *total* reconstructed mass map is zero, so the analysis is restricted to half of the grid size.

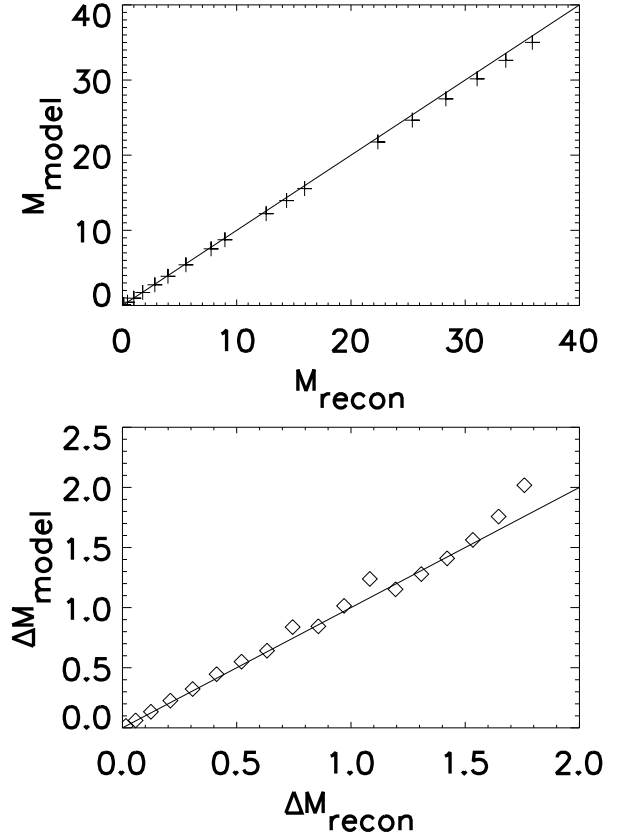


Figure 3. The crosses on the first plot show the total mass M_{recon} measured in the simulation for different sizes of the aperture S and for the gaussian smoothing case. For the smallest aperture, which corresponds to the pixel size, $M_{recon} \simeq 0.1$. The largest aperture has a width of 18 pixels (about half of the simulation grid), and $M_{recon} \simeq 36$. The vertical axis displays the smoothed mass distribution of the model. The solid line shows a perfect correlation between M_{recon} and M_{model} . The reconstructed mass is overestimated by less than 2%. The diamonds on the second plot show the standard deviation of the mass for different aperture sizes. The horizontal axis corresponds to the measure of ΔM in the simulation, and the vertical axis is the prediction from Eq.(27). The solid line shows a perfect correlation between ΔM_{recon} and ΔM_{model} .

4.2 Effect of the mass-sheet degeneracy

For each reconstruction, the mass map is rescaled in such a way that the total mass measured on each realization is the same as the total mass in the smoothed lens model.

Such a rescaling changes the amplitude of the noise correlation function: if K_{real} is the real mass distribution, the reconstructed mass map K_{recon} is related to the real mass map by $K_{recon} = 1 - \lambda + \lambda K_{real} + N$, where λ is the mass-sheet degeneracy constant and N the noise term. Therefore the estimate of the real mass distribution is $K_{real} + \frac{N}{\lambda}$ where the noise scales now with λ^{-1} . It follows that the noise correlation function of the reconstructed mass map scales as λ^{-2} , according to Eq.(22). In that sense, error bars of K depend on the absolute value of K , although it is not explicit in Eq.(22), and this rescaling has to be taken into account when the amplitude of the noise correlation function is calculated.

4.3 Correlation functions

Figure 2 is a comparison of the measured noise correlation function to the predictions (Eq.(26) for the Gaussian smoothing and Eq.(29) for the top-hat smoothing). The dashed lines correspond to the Gaussian smoothing and the solid lines to the top-hat. The thin lines display the measurement in the simulations, they are in very good agreement with the predictions (thick lines). As expected, the correlation drops to zero for the top-hat window when $|\mathbf{r}| = 2\sigma$. The amplitude of the correlation function is twice the value for the top-hat compared to the Gaussian, as expected from Eq.(26) and (29).

4.4 Mass error bar

The average variance of the mass in an area $\mathcal{S} = 2a \times 2a$ is given by Eq.(23). This quantity does not depend on the signal according to our hypothesis. As shown by the plot at the top of Figure 3 the reconstructed mass is an unbiased estimate of the smoothed initial mass distribution (as expected from Eq.(15)). The reconstructed mass is slightly overestimated by less than 2%.

The diamonds on the bottom plot of Figure 3 shows the error Eq.(27) (ΔM_{model}), compared to the measured error on the simulations (ΔM_{recon}). The different diamonds correspond to different sizes of the smoothing area $\mathcal{S} = 2a \times 2a$ ($\Delta_{recon} \simeq 0.1$ for $2a = 1 \text{ pixel}$, and $\Delta_{recon} \simeq 1.5$ for $2a = 18 \text{ pixels}$). The agreement between the model and the measurement in the simulation is good, however, there is a tendency of the measured variance to be slightly underestimated for $2a \simeq 18$. This is a consequence of the normalisation of the mass maps (the total measured mass is explicitly set to the total mass in the model, so the measured variance is zero at the scale $2a = 42 \text{ pixels}$). The errors are independent of the position of \mathcal{S} on the grid, which confirms that signal and noise are uncorrelated to a very good approximation.

5 SIGNIFICANCE OF PEAKS

The reconstructed mass maps are generally very noisy, even with an arcmin scale smoothing, and it is difficult to disentangle real and noise over-densities. Figure 4 is an example of mass reconstruction when no lenses are present. It shows fluctuations which can be misinterpreted as real mass over-densities. It makes particularly difficult the selection of objects based on their mass, and in practice we use a high signal-to-noise cut-off to be sure to minimize the number of spurious peak detections (Schneider 1996, Reblinsky & Bartelmann 1999). According to Section 3.2, the noise in the reconstructed mass maps is equivalent to the convolution of a point process by a smoothing aperture, therefore it is expected to approach a Gaussian field distribution when the aperture is larger than the mean angular distance between galaxies. The nice feature of Gaussian fields is that they are fairly well understood (Bardeen et al. 1986), in particular the statistics of two-dimensional Gaussian random fields as been worked out in detail for the case of cosmic microwave background fluctuations (Bond & Efstathiou 1987, hereafter BE).

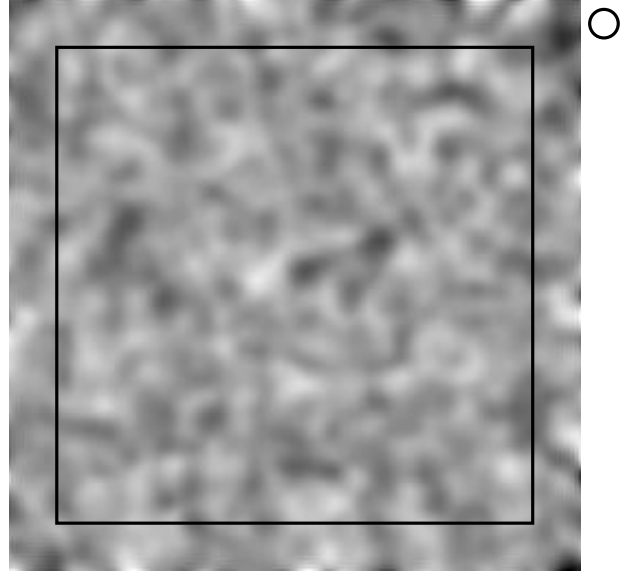


Figure 4. One realization of mass reconstruction with no lensing inside. All the structures here are noise. The high density peak at the center is a 3σ fluctuation ($\kappa = 0.081$). The dark square represents the subfield where the statistical analysis is done (in order to avoid the noisier boundaries), and the circle at the top right has a $0.66'$ diameter, representing the size of the Gaussian window width. The field is $14' \times 14'$ large, with a number density of 30 galaxies per arcmin^2 . The intrinsic ellipticity dispersion is $\sigma_e = 0.17$.

I propose here to use the statistics of 2D Gaussian random fields as a tool to analyze the weak lensing mass maps. With this formalism, the significance and the morphology of the structures can be evaluated analytically and compared with the data, which is a way to go beyond the detection of dark halos done in Schneider (1996), based on a local signal-to-noise analysis.

5.1 Height of peaks

Let $N(\boldsymbol{\theta})$, the noise part of Eq.(16), be a 2D Gaussian random field. The field and its derivatives ($\mathbf{V} = [N, N', N'']$) are distributed as

$$\mathcal{P}(\mathbf{V}) = \frac{1}{(2\pi)^{3/2} |\mathbf{M}|^{1/2}} \exp\left(-\frac{1}{2} \mathbf{V}^t \cdot \mathbf{M}^{-1} \cdot \mathbf{V}\right), \quad (34)$$

where the correlation matrix \mathbf{M} is

$$\mathbf{M} = \begin{pmatrix} \sigma_0^2 & 0 & -\sigma_1^2 \\ 0 & \sigma_1^2 & 0 \\ -\sigma_1^2 & 0 & \sigma_2^2 \end{pmatrix}. \quad (35)$$

The σ_n are determined from the noise power spectrum

$$\sigma_n^2 = \int \frac{k dk}{2\pi} k^{2n} \langle |\tilde{N}(k)|^2 \rangle, \quad (36)$$

which can be calculated from the two-point correlation function:

$$\sigma_n^2 = (-1)^n \frac{2\pi}{\int_0^{2\pi} d\phi \cos^{2n}(\phi)} \left(\frac{d^{2n} \langle N(\boldsymbol{\theta}) N(\boldsymbol{\theta} + \mathbf{r}) \rangle}{dr^{2n}} \right)_{r=0}. \quad (37)$$

The zero-lag value of the correlation function is $\sigma_0^2 = \langle N(0) N(0) \rangle$. The marginalisation of Eq.(34) over the field

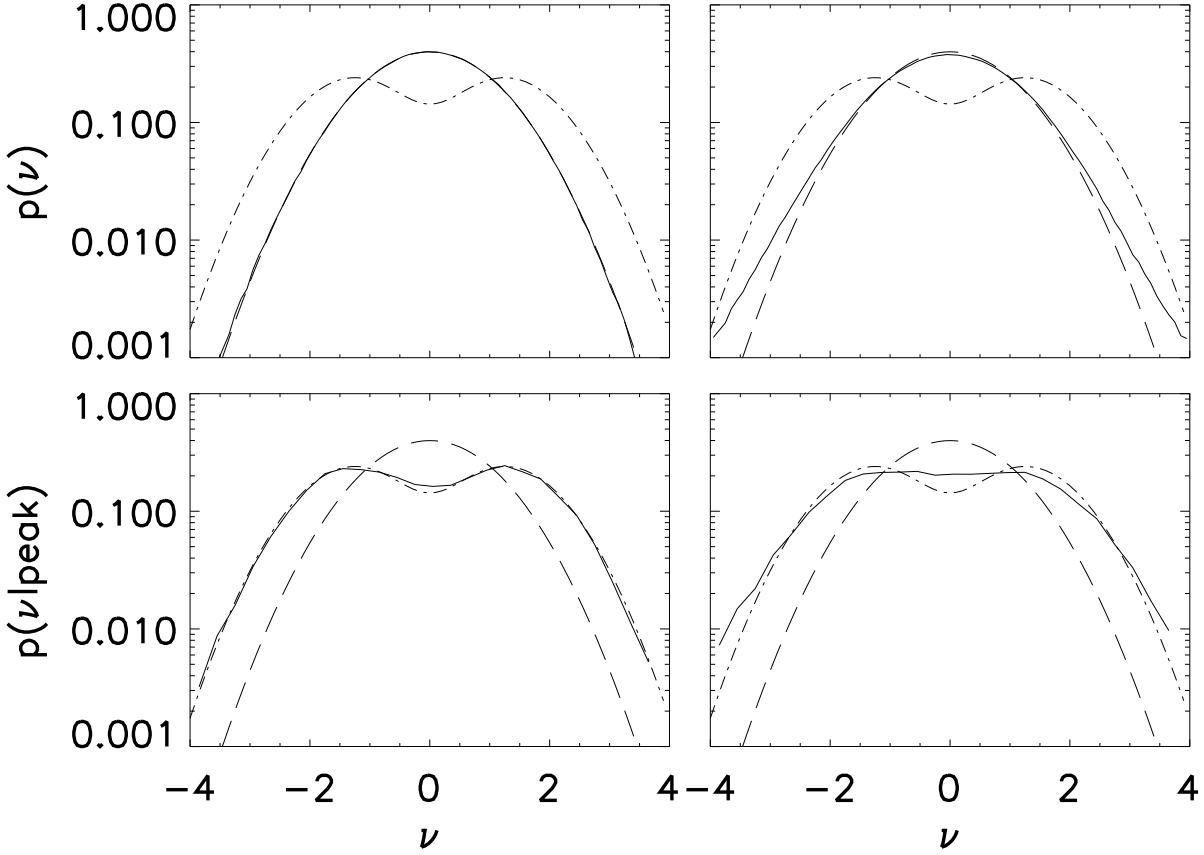


Figure 5. In all the plots, the dashed curve is the probability distribution function that a randomly chosen pixel has a value $\nu\sigma_0$ (Eq.(38)), and the dot-dashed line is the expected probability distribution of peaks extrema, according to Eq.(40). The solid line in the top-left plot is the measured distribution of all the pixels in the 90 simulations (described in Section 5.1.) limited to the sub-field defined by the dark square as explained in Figure 4. The solid line in the bottom-left plot shows the distribution of extrema measured in the same sub-fields of the simulations. In the top-right plot, the solid line is the measured distribution of the value of all the pixels in the whole field, and in the bottom-right plot, the solid line is distribution of extrema measured in the whole field.

derivatives gives the -Gaussian- probability that any pixel randomly chosen in the field has the value $\nu\sigma_0$:

$$\mathcal{P}(\nu) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\nu^2}{2}\right). \quad (38)$$

In order to describe the probability distribution function of peaks, the following parameters are introduced: γ and θ_* are respectively the width of the power spectrum and the characteristic angular scale of the field:

$$\gamma = \frac{\sigma_1^2}{\sigma_0\sigma_2}; \quad \theta_* = \sqrt{2}\frac{\sigma_1}{\sigma_2}. \quad (39)$$

A peak occurs, by definition, where the gradient of the field vanishes ($\nabla_{\theta}N = 0$), and it has a given height $\nu\sigma_0$. Its probability distribution function $\mathcal{P}(\nu|peak)$ can be calculated from the number density of minima and maxima $\mathcal{N}_{max}(\nu)$ and $\mathcal{N}_{min}(\nu)$ (see BE for the details):

$$\mathcal{P}(\nu|peak) = \frac{\mathcal{N}_{max}(\nu) + \mathcal{N}_{min}(\nu)}{\int_{-\infty}^{\infty} d\nu (\mathcal{N}_{max}(\nu) + \mathcal{N}_{min}(\nu))}, \quad (40)$$

where

$$\mathcal{N}_{max}(\nu) = \mathcal{N}_{min}(-\nu) = \frac{1}{2\pi\theta_*} \exp\left(-\nu^2/2\right) \frac{G(\gamma, \gamma\nu)}{\sqrt{2\pi}}. \quad (41)$$

The function $G(\gamma, \bar{x})$ depends on the variable $b^2 = 2(1-\gamma^2)$:

$$G(\gamma, \bar{x}) = \frac{1}{2} \left(\bar{x}^2 + \frac{b^2}{2} - 1 \right) \operatorname{erfc}\left(-\frac{\bar{x}}{b}\right) + \frac{\bar{x}b}{2\sqrt{\pi}} e^{-\frac{\bar{x}^2}{b^2}} + \frac{1}{2(1+b^2)^{1/2}} e^{-\frac{\bar{x}^2}{1+b^2}} \operatorname{erfc}\left(-\frac{\bar{x}}{b\sqrt{1+b^2}}\right). \quad (42)$$

The denominator of Eq.(40) has an analytical solution which can be derived straightforwardly with a formal calculator, but it is not worth to give its expression here.

For a Gaussian smoothing (Eq.(25)) the spectral parameters take the values $\gamma = 1/\sqrt{2}$ and $\theta_* = \sigma/\sqrt{2}$, and therefore $\mathcal{P}(\nu|peak)$ only depends on ν . The denominator of Eq.(40) is:

$$\int_{-\infty}^{\infty} d\nu (\mathcal{N}_{max}(\nu) + \mathcal{N}_{min}(\nu)) = \frac{1}{(2\pi)^{3/2}\theta_*} \frac{4\sqrt{3}\pi - 9}{9\sqrt{2\pi}} \quad (43)$$

It means that once the mass map is normalised by σ_0 the probability distribution of peaks is always the same, independently of the smoothing scale, the number density of background galaxies, and the ellipticity dispersion σ_e^2 .

The analytical predictions of Eq.(38) and Eq.(40) are now compared to numerical simulations. Ninety mass recon-

structions have been performed, with no lensing signal inside. The total field of view for each realization is 4100×4100 pixels (equivalent to a $14' \times 14'$ area at the Canada France Hawaii Telescope), and it contains about 30 gal/arcmin^2 . The intrinsic ellipticity dispersion is $\sigma_e = 0.17$ (using the distribution Eq.(8)), and the smoothing aperture is given by Eq.(25) with $\sigma = 0.33'$, which corresponds to an average of 10 galaxies per aperture. The "mass" field is reconstructed using the maximum-likelihood algorithm. Since the aperture is truncated at the edge of the field, only a sub-set of pixels in the simulations are used, those enclosed in the square area defined on Figure 4. This Figure also shows the scale of the smoothing window compared to the whole field.

Figure 5 shows the measured probability distribution functions in the field and for peaks, for the whole field of the simulations, and for the sub-field as described in Figure 4. The measurements are represented by the solid lines, while the dashed and dot-dashed curves show respectively the unconstrained (Eq.(38)) and the constrained (Eq.(40)) distributions. For the centered sub-field it is remarkable that the peak distribution prediction fits so well the measurements. It shows that the non-linear ML reconstruction give a noise comparable to the noise in weak lensing approximation, confirming our previous results. However it is clear that when the measurement is done in the whole field, the statistic is not gaussian. This is an effect caused by the boundary of the reconstruction: at the edge of the field, the smoothing aperture is truncated and it contains less galaxies than in the central part of the simulations. Therefore the peak analysis has to be restricted to the area where the aperture is not too close to the edge. Two other sets of 20 simulations each, one with $\sigma_e = 0.4$ and one with a uniform distribution of ellipticities between -0.2 and 0.2 (instead of Eq.(8)) give the same result, which shows that the Gaussian behavior of the noise does not depend on the details of the ellipticity distribution. The distribution of peaks is not Gaussian, as expected for Gaussian fields. It shows that simple signal-to-noise analysis based on the local variance with Gaussian probabilities is wrong. For instance a value at 3σ in the field, with $\mathcal{P}(\nu) = 0.004$ corresponds to $\mathcal{P}(\nu|peak) = 0.03$ for a peak, almost a factor of ten higher.

Because of the discreteness of the spatial distribution of the galaxies a Poisson-like noise is expected rather than a Gaussian noise. The fact that the noise is not Poisson is a consequence of the central limit theorem which makes the distribution more Gaussian because there is an average of 10 galaxies per aperture, which is much larger than 1.

The peak statistic described here is very easy to implement and it requires only the calculation of σ_0 . In other words, only the galaxy number density, the variance of their ellipticities and the shape of the smoothing window are needed.

5.2 Profile of peaks

It is useful to add the morphological information to the peak analysis in order to constrain the mass distribution better. Let $\mathcal{P}(shape|\nu)$ be the probability that a peak of height $\nu\sigma_0$ has a given shape, and $\mathcal{P}(shape, \nu)$ the global peak probability written as:

$$\mathcal{P}(shape, \nu) = \mathcal{P}(\nu|peak)\mathcal{P}(shape|\nu) \quad (44)$$

BE give the analytical formulae describing the mean profile of a peak and its dispersion. This calculation depends on the averaged curvature of a peak of height ν . The curvature x is defined from the eigenvalues (λ_1, λ_2) of the second derivative matrix of the noise $\partial_{ij}N$:

$$x = (\lambda_1 + \lambda_2)/\sigma_2. \quad (45)$$

Following BE, all distances and derivatives are calculated with respect to the reduced coordinates $\omega = \mathbf{r}/\theta_*$. The reduced correlation function is defined as $\psi(\omega) = \langle N(\theta)N(\theta + \omega) \rangle / \langle N(0)N(0) \rangle$. The profile around a peak, averaged over the curvature and the ellipticity, is Gaussian distributed with a mean $\langle N(\omega)|\nu \rangle$ and a dispersion $\langle N^2(\omega)|\nu \rangle$ (Eqs. A2.9 and A2.8 of BE):

$$\begin{aligned} \langle N(\omega)|\nu \rangle / \sigma_0 &= \frac{\nu}{1 - \gamma^2} (\psi + \Delta\psi/2) - \frac{\langle x|\nu \rangle / \gamma}{1 - \gamma^2} (\gamma^2 \psi + \Delta\psi/2) \\ \langle N^2(\omega)|\nu \rangle / \sigma_0^2 &= 1 - \frac{1}{1 - \gamma^2} \psi^2 - \frac{1}{1 - \gamma^2} \psi \Delta\psi \\ &\quad - \frac{1}{\gamma^2(1 - \gamma^2)} (\Delta\psi/2)^2 - \frac{1}{\gamma^2} (\psi')^2 \\ &\quad - \frac{1}{2\gamma^2} [\omega(\psi'/\omega)]^2. \end{aligned} \quad (46)$$

For 2D Gaussian fields, the conditional mean value of the average curvature $\langle x|\nu \rangle$ can be calculated analytically:

$$\langle x|\nu \rangle = \gamma\nu + \frac{H(\gamma, \gamma\nu)}{G(\gamma, \gamma\nu)}. \quad (47)$$

The function $H(\gamma, \bar{x})/G(\gamma, \bar{x})$ tends to zero for high peaks ($\nu \rightarrow \infty$). $G(\gamma, \bar{x})$ was given in Eq.(42) and $H(\gamma, \bar{x})$ is:

$$\begin{aligned} H(\gamma, \bar{x}) &= \frac{\bar{x}b^2}{2} \text{erfc}\left(-\frac{\bar{x}}{b}\right) + \frac{b^5}{2\sqrt{\pi}(1+b^2)} e^{-\frac{\bar{x}^2}{b^2}} \\ &\quad - \frac{\bar{x}b^2}{2(1+b^2)^{3/2}} e^{-\frac{\bar{x}^2}{1+b^2}} \text{erfc}\left(-\frac{\bar{x}}{b\sqrt{1+b^2}}\right) \end{aligned} \quad (48)$$

The quantity χ^2 quantifies how any observed peak profile of height $\nu\sigma_0$ $N_{obs}(\theta, \nu)$ is different from its expectation value:

$$\chi^2 = \sum_{\omega} \frac{(N_{obs}(\omega, \nu) - \langle N(\omega)|\nu \rangle)^2}{\langle N^2(\omega)|\nu \rangle}. \quad (49)$$

This follows a χ^2 statistic with the number of degrees of freedom n given by the number of positions ω involved in the summation. Thus the probability $\mathcal{P}(shape|\nu)$ that a peak of height $\nu\sigma_0$ having a definite shape is a noise fluctuation is,

$$\mathcal{P}(shape|\nu) = \frac{1}{2\Gamma(\frac{n}{2})} \int_{\chi^2}^{\infty} \left(\frac{t}{2}\right)^{\frac{n}{2}-1} e^{-\frac{t}{2}}. \quad (50)$$

The probability $\mathcal{P}(shape, \nu)$ that a peak of height $\nu\sigma_0$ with a given shape is a noise fluctuation is then given by Eq.(44), which contains the probabilities (40) and (50). Figure 6 is an example of mean profile and its dispersion derived from Eqs.(46).

6 CONCLUSION

It is possible to understand the noise properties of reconstructed mass maps with rather simple arguments. The noise

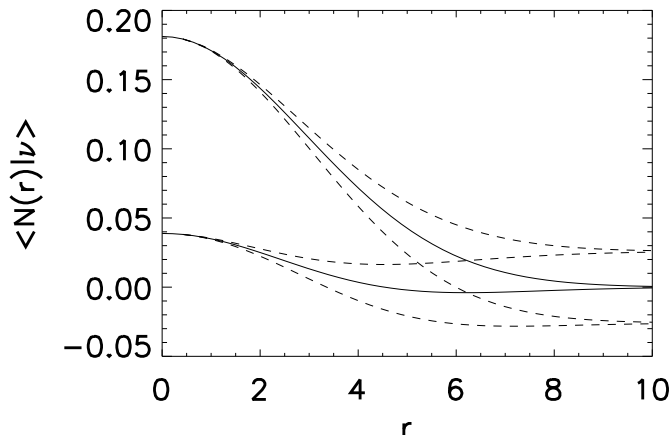


Figure 6. Mean profile around a peak of height $7\sigma_0$ (top curve) and $1.5\sigma_0$ (bottom curve) with the noise correlation function given by Eq.(26). The parameters are $\sigma = 3$, $n_g = 0.6$ galaxies/pixel², and $\sigma_\epsilon = 0.12$. The dashed curves represent the $\pm 1\sigma$ deviations.

in ML mass reconstruction is very close to a 2D Gaussian random field, which has several implications for the analysis of mass maps: for cluster lensing, the "blurred" area around cluster of galaxies where the identification of mass clumps and substructures is uncertain can in fact be described accurately. It is possible to give the probability that the fluctuations are real mass-over-densities, and to give their mass with an error bar. For large scale surveys, small mass concentrations (not visible individually) can be detected statistically as a deviation from the expected noise Gaussian statistic. Groups and small clusters should show up in the statistical distribution of peaks. Preliminary results using ray-tracing simulations (Jain, Seljak, White 1999, Jain private communication) already confirmed this, but the detailed analysis of peak statistic in realistic cosmological models is left for another work.

The noise autocorrelation function and the mass error bars were derived and compared to numerical simulations of non-critical cluster lenses. The agreement between the predictions and the simulations proves that the noise, caused by the intrinsic ellipticities and the discrete distribution of the galaxies, is nearly uncorrelated with the lensing signal despite the non-linear lensing equation. Consequently, the weak lensing approximation applies, which simplifies considerably the noise analysis. It is worth to remind that this is true for non-critical clusters only. However it is reasonable to think that the smearing of the signal due to the smoothing implies that the approximation still works for noise analysis in critical clusters (if the smoothed mass distribution becomes non-critical).

A BBKS-type approach was used to predict the statistics of noise peaks in mass maps. Again the numerical simulations are accurately fitted by these predictions. The fact that the noise approaches a Gaussian distribution is due to the large mean number of galaxies in each aperture (about 10). There is certainly a limitation of the method used here: if the smoothing window is too small or too "fuzzy" it should

deviate from a Gaussian statistic. This departure from Gaussianity was observed in the S-statistic of Schneider (1996) because he used a compensated filter. A work worth to be done is to investigate the limits of the method, and eventually to consider the Poisson behavior of counts in small cells, as in Szapudi & Colombi (1996). However, in the context discussed here, it is reliable to use Gaussian statistics.

I discussed so far only the ML mass reconstruction method, and compared it to the analytical KS. Can this analysis be used for other mass reconstruction methods? There are some numerical indications that the curl-free line-integral approaches are close to the KS reconstruction, (Seitz & Schneider 1996, Lombardi & Bertin, 1998b) and therefore close to the ML as well. For the newly developed regularization methods (Seitz et al. 1998, Bridle et al. 1998) it probably not the case. The noise is reduced in these methods, in an optimal way that real structures should still emerge from the noise. The main advantage of the regularization methods is their adaptive spatial resolution depending on the local signal-to-noise. However the regularization parameter is arbitrary and can potentially kill the small scale fluctuations that we are looking for, so the analysis done here should not apply. It is therefore valuable to make a study of the noise properties for all the reconstruction methods in order to search for a better use of the noisy regions of the mass maps and to compare their ability to detect the mass overdensities.

7 ACKNOWLEDGMENTS

I thank R. Scoccimarro, Y. Mellier and T. Erben for a careful reading of the manuscript.

REFERENCES

- Bardeen, J. M., Bond, J.R., Kaiser, N., Szalay, A.S., 1986, ApJ, 304, 15
- Bartelmann, M., Narayan, R., Seitz, S., Schneider P., 1996, ApJ, 464, L115
- Bond, J.R., Efstathiou, G., 1987, MNRAS, 226, 655 (BE)
- Bridle, S.L., Hobson, M.P., Lasenby, A.N., Saunders, R., 1998, MNRAS, 299, 895
- Connolly, A.J., Szalay, A.S., Brunner, R.J., 1998, ApJ, 499, L125
- Erben, T et al., astro-ph/9907134
- Hattori, M., 1997, "Cosmic Chemical Evolution", IAU 187, Kyoto, Japan
- Jain, B. Seljak, U., White, S., astro-ph/9901191
- Kaiser, N. 1995, ApJ, 439, L1
- Kaiser, N. 1998, ApJ 498, 26
- Kaiser, N., Squires, G., 1993, ApJ 404, 441
- Kaiser, N. et al., astro-ph/9809268
- Le Fèvre, O., Hudon, D., Lilly, S.J., Crampton, D., Hammer, F., Tresse, L., 1996, ApJ, 461, 534
- Lombardi, M., Bertin, G., 1998a, A& A, 335, 1
- Lombardi, M., Bertin, G., 1998b, A& A, 330, 791
- Lombardi, M., Bertin, G., 1999, A& A, 348, 38
- Oukbir, J., Blanchard, A., 1997, A& A, 317, 1
- Peebles, P.J.E., 1980, *The large scale structure of the Universe*, Princeton
- Reblinsky, K., Bartelmann, M., 1999, A& A, 345, 1
- Reblinsky, K., Kruse, G., Jain, B., Schneider, P., astro-ph/9907250
- Schramm, T., Kayser, R., 1995, A& A, 299, 1

Schneider, P., 1996, MNRAS, 283, 837
Schneider, P., Seitz, C., 1995, A& A, 294, 411
Seitz, S., Schneider, P., 1996, A& A, 305, 383
Seitz, C., Schneider, P., 1997, A& A, 318, 687
Seitz, S., Schneider, P., Bartelmann, M. 1998, A& A, 337, 325
Squires, G., Kaiser N., 1996, ApJ 473, 65
Szapudi, I., Colombi, S., 1996, ApJ, 470, 131
Van Waerbeke, L., Bernardeau, F., Mellier, Y., 1999, A& A,
342, 15
Wilson, G., Cole, S., Frenk, C.S., 1996, MNRAS, 280, 199